

Figure 1. Amino acid alignment of soybean CDPKa (SEQ ID NO: 2) and CDPKb (SEQ ID NO: 4), maize CDPK (L27484, SEQ ID NO: 23) and Arabidopsis CDPK (U20388, SEQ ID NO: 24)

```

U20388      MANKPRTRWVLPYKTKNVEDNYFLGQVLGQGQFGTTFLCTHKQTGQKLACKSIKPKRLLC
CDPKb      -----
L27484      -----MEDVRATYSMGKELGRGQFGVTHLCTHRTSGEKLACKTIAKRKLAA
CDPKa      -----MTT

U20388      QEDYDDVLREIQIMHHLSEYPNVVRIESAYEDTKNVHLMELCEGGELFDRIIVKRG-HYS
CDPKb      -----MELCAGGELFDRIIQRG-HYT
L27484      REDVDDVRREVQIMHHLSGQPNVVGRLGAYEDKQSVHLMELCAGGELFDRIIARG-QYT
CDPKa      AIAIEDVRREVKILRALTGHNLVQFYEAYEDDDNVYIVMELCKGGELLDRIILSRGGKYS
                ****  *****:***:  **  *:

U20388      EREAAKLIKTIIVGVVEACHSLGVVHRDLKPENFLFSSSDEDASLKSTDFGLSVFCTPGEA
CDPKb      ERQAAKLTKTIVGVVEACHSLGVMHRDLKPENFLFVNQHEDSLKTIIDFGLSVFFKPGDI
L27484      ERGAAELLRAIVQIVHTCHSMGMHRDIKPENFLLSKDEDAPLKATDFGLSVFFKEGEL
CDPKa      EEDARVVMIQILSVVAFCHLQGVVHRDLKPENFLFTSKDDKSTLKAIDFGLSDYVKPDER
                * . * : * : * * ** **:*****: ..... **: ***** : . . :

U20388      FSELVGSAYYVAPEVLHKHYGPECDVWSAGVILYILLCGFPFWESEIGIFRKILQGKL
CDPKb      FNDVVGSPYYVAPDVLKRKYGPEADVWSAGVILYILLSGVPPFWEENEQGI FEQVLRGDL
L27484      LRDIVGSAYYIAPEVLKRKYGPEADIWSVGVMLYIFLAGVPPFWEENENGIFTAILRGQL
CDPKa      LNDIVGSAYYVAPEVLHRSYGTADMWSIGVIAIYILLCGSRPFWARTESGIFRAVLKADP
                : : :***.***:***:***: **.*.*** ** : **.* * **.*. * * * :*:..

U20388      EFEINPWPSISESAKDLIKKMLESNPKKRLTAHQVLCHPWIVDDKVAPDKPLDCAVVSRL
CDPKb      DFSSDPWPSISESAKDLVRKMLVRDPRRRLTAHQVLCHPWIQVDGVAPDKPLDSAVLSRL
L27484      DLSSEPWPHPISPGAKDLVKMMLNINPKERLTAHQVLNHPWIKEDGDAPDTPLDNVVLDRL
CDPKa      SFDEAPWPSLSVDAKDFVKRLNKDYRKRLTAAQALSHFWLVNHHDDMRIPLDMIHKL
                :.. *** : * .*****: * : :***** *.* ***: . *** : . :

U20388      KKFSAMNKLKKMALRVIAERLSEEEIGGLKELFKMIDTDKSGTITFEELKDSMRVSGEL
CDPKb      KQFSAMNKLKKMALIIAESLSEEEIAGLKEMFKMIDADNSGQITFEELKAGLKRVGANL
L27484      KQFRAMNQFKKALRIIAGCLSEEEITGLKEMFKNIDKDNSGTITLDELKHGLAKHGPKL
CDPKa      KAYICSSSLRKSALRALAKTLTVAQLAYLRDQFTLLGPNKSGLISMQNFKTAVLRSSDA
                * : . . . : * * * : * * : : * : * . . : * * * : : * : . . .

U20388      M-ESEIQELLRAADVDESGTIDYGEFLAATIHNLKLE----REENLVAAFSFFDKDASGY
CDPKb      K-ESEIYDLMQAADVDSNGTIDYGEFLAATLHRNKIE----REDNLFAAFSYFDKDGSGY
L27484      S-DSEMEKLMEAADADGNGLIDYDEFVTATVHMNKLD----REEHLYTAFQYFDKDNSGY
CDPKa      SKDSRVLDYVSMVSSIQYRKLDFFEECAAISVHQLEGMETWEQHARHAYELFKKEGNRP
                :*.: . : . . :*: ** :*: : :*: *.: *.: *.*: .

U20388      ITIEELQQAWKEFGINDS-NLDEMIKDIDQDNDGQIDYGEFVAMMRKNGTGGGIGRRTM
CDPKb      ITQEELQQACDEFGIKDV-RLEEIIKEIDEDNDGRIDYNEFVAMMQKGN--LPAVGKKGL
L27484      ITKEELEHALKEQGLYDADKIKDIISDADSDNDGRIDYSEFVAMMRKGT---AGAEPMNI
CDPKa      IMIEELAS---ELGLSPSPVHVVLQDWIRHSDGKLSFLGFVRLH-----GVSSRAF
                * *** * * : . : : : . ** : : * * : : . : :

U20388      RNSLNFGTTLPDESMNV
CDPKb      EN--SFSVKFR-EALKL
L27484      KK--RRDIVL-----
CDPKa      QKA-----
                .:

```

Figure 2: 5' Flanking Region of the soybean AMPD gene (SEQ ID NO: 5) showing the TATA box and first ATG codon in bold face type.

GGCTTGTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGACATTCTGAGGAAAGAGCC  
AGAGCATGAAACTTTTACAAGATTGAGAATAACTCCTCTTGGTATGAACTCTGACATGTTACTCTTCCTCATTCT  
CTTGAGTGCAAAGAAATTGCTTTGAATGGGTTTTTCATGGGATGTRTAGAATCCCCCTCCTTCARACACTTGTTT  
GATTACAATATTTGCATATCTGGAAAATTTATAGTAACTATAATTGTTATATTTTTTGTGTTAGAGGCTCCGTCA  
CCTGATGAAATTGAAGCTTATGTGGTTCTGCAAGAATGCCTTGAAATGAGAAAAAGATATGTTTTTAGAGAAGCT  
GTTGCTCCGTGGGATAAAGAAGTTATATCCGACCCCAGCACACCCAAGCCTAACCCAGATCCATTTTTATACATT  
CCTGAAGGAAATTCTGATGTGAGTTTTTTTCTCCCACCACTGAAAACCTGGATCTCCCTATGTTATATGCTGTGA  
TTAAATTAGTTGTATTTCTCTTGTGCTACAGCATTATTTTGAAATGCAAGATGGGGTTATTTCGTGTATATCCAGA  
TAGAGATGGTAAGTAACAGGAATTTGTTATTGATAAACTTGGTAATTTTATCACCCTCAAGGTGGGATCTTATG  
CATGCTTGCCATTACTTGCAGCAAAAGAAGAGCTTTTTCCTGTAGCTGATGCAACTACATTTYTCACYGATCTT  
CATCACTTACTTCGAGTCATAGCAGCAGGGAAT**TATA**AGAACTTTATGCCATCATAGGCTCAATCTTCTAGAACAA  
GTACATCTCTAATTTACTGAAACAACTGCAGCCTATGCTTGTATTTTAATTACATACAAGAATCAATTGTTGTT  
TGACAATTTCTGTATTTTTTCAGAAATTCAATCTTCATTTG**ATG**

Figure 3: Amino acid alignment of soybean NRTF1a (SEQ ID NO: 7), NRTF1b (SEQ ID NO: 9), with two *Arabidopsis thaliana* AP2 proteins, AJ001911 (SEQ ID NO: 31) and AF003096 (SEQ ID NO: 32), with the conserved AP2-domain indicated by underlining.

```

NRTF1b      MCGGAIISDFIGVK---RGRNLAAQELWSELDP--FSDFLGFDTTNSK--NQPPLOKIPD
NRTF1a      MCGGAIISDFIGVK---RGRNLAAQELWSELDP--FSDLLGFDTTTTTTTNQPP---LPD
AJ001911    MCGGAIISDYAPLVTKAKGRKLTAEELWSELDA SAADDFWGFYSTSKL---H-----PT
AF003096    MCGGAIISDYAPLVTKAKGRKLTAEELWSELDA SAADDFWGFYSTSKL---H-----PT
*****:   :   :*:*:*:*****.   .*: ** :*..   :   *

AP2-2      KKVVSSEKKKKSVVGAEEKKSDSGRARKNVYRGIRQRPWGKWAAEIRDPHKGVRVWLGT
AP2-1      KKVVSSEKKKKKSVAEKK-S-GGRARKNVYRGIRQRPWGKWAAEIRDPHKGVRVWLGT
AJ001911    NQVNVKEEEAVKKEQATEPGKR---RKRKNVYRGIRKRPWGKWAAEIRDPKRGVRVWLGT
AF003096    NQVNVK-EEAVKKEQATEPGKR---RKRKNVYRGIRKRPWGKWAAEIRDPKRGVRVWLGT
.: * . *: * . :.*      * *****:*****:*****

AP2-2      FPTAEAAQAYDDAAIRIRGDKAKLNFPATTISAAA--APPPS-----KKQRCL
AP2-1      FPTAEAAARAYDDAAKIRIRGDKAKLNFPAT-----APPPS-----KKQRCL
AJ001911    FNTAEAAAMAYDVA AKQIRGEKAKLNFPDL DHPS---TPPPSSTSLRLSDQPPAKKVCV
AF003096    FNTAEAAAMAYDVA AKQIRGDKAKLNFPDLHHP PPPNYTPPPS--SPRSTDQPPAKKVCV
* ***** ** * :*:*:*****      :****      *: *:

AP2-2      SPDIITESSSSSSHSTTGSTGE---SG---GGNDEL DLKQ----IEWFLGLENELPVSN
AP2-1      SPDTTTEQSSSS--QSTTGSTGSP-PSAAFHGGGDEL DLKQ----LERFLGLD-----
AJ001911    VSQSE--LAQPSFPVECVGFGKGEEFQNL MYGFEPDYDLKQQISSLESFLELDGTAEQP
AF003096    VSQSESELSQPSFPVE CIGFGNGDEFQNL SYGFEPDYDLKQQISSLESFLELDGNTAEQP
.:      :..* . * . * : **** :* ** *:

AP2-2      NIGAEDNMDDLWMLDDVVVPNRHLIY
AP2-1      NMGAEDNMDDLWMLDDVVVPNRHLIY
AJ001911    SQLDESVCDDVMWMLDDVIASYE----
AF003096    SQLDES VSEVDMWMLDDVIASYE----
.   *      *:*****:..

```

Amino acid alignment of soybean NRTF1a (SEQ ID NO: 7), NRTF1b (SEQ ID NO: 9), NRTF1c (SEQ ID NO: 11) and NRTF1d (SEQ ID NO: 13), with the conserved AP2-domain indicated by underlining.

```

NRTF1a- MCGGAIISDFI--GVKRGGRN-LAAQELWSELDP-FSDLLGFDTTTTTT---TNQPP---
NRTF1b MCGGAIISDFI--GVKRGGRN-LAAQELWSELDP-FSDFLGFDTTNSK-----NQPP LQKI
NRTF1c_ MVSATVDS DFA--FLESVQQYLLGHDSINLMSE-THQAASHDPFSDP-----NKCD----
NRTF1d_ MVSATVDS DFA--FLESVQQYLLGHDSINLMSE-THQAASHDPFSDP-----NKCD----
      *  . . : : ***          . : . .      :      .      *          : .

NRTF1a- PDKKVVS SCEEKKKK-----KVS SAEKK-S-G-----G-----RARKNVYRGIRQ
NRTF1b PDKKVVS SCEEKKKK-----SVVGA EKKKSDS-----G-----RARKNVYRGIRQ
NRTF1c_ GDSGNI AFRSEDAT-----AVVARDHA-----PPTWKHYRGVRR
NRTF1d_ GDSGNI AFRSEDAT-----AVVARDHA-----PQTWKHYRGVRR
      * .          . . .          . . :          : : *** : :

NRTF1a- RPWGKWA AEIRDPHK-GVRVWLGT FPTAEEAARAYDDAAK RIRGD KAKLNF PAT-----
NRTF1b RPWGKWA AEIRDPHK-GVRVWLGT FPTAEEAAQAYDDAAIRIRGD KAKLNF PAT TISAAA
NRTF1c_ RPWGKFA AEIRDPKKN GARVWLGT YDTEEKAALAYDKAAFKMRGQKAKLNF PHL-----
NRTF1d_ RPWGKFA AEIRDPKKN GARVWLGT YDTEEKAALAYDKAAFKMRGQKAKLNF PHL-----
      ***** : ***** : * * . ***** : * * . ** * * * * : : * . *****

NRTF1a- APPPSKKQRCLSPDTTTEQSS--SQSTTGSTGSPPSAAFHG GDEL DLKQLERFLGLD-
NRTF1b APPPSKKQRCLSPDIITESSSSSSHSTTGSTGE--SG---GGNDEL DLKQIEWFLGLE N
NRTF1c_ --IDSDNSELSEPVMMTTSKR--SLLEISSPSS-----SCSDDSSESQ--GTKRRK-
NRTF1d_ --IDSDNSELSEPVMMTTSKR--SLLEISSPSS-----SYSDDSSESQ--GTKRRK-
      * . :      **          * .      *          * .      .          . . :      .

NRTF1a- -----NMGA EWDNMDDLW---MLDDVVVPNRHLIY-----
NRTF1b ELPVSNNIGA EWDNMDDLW---MLDDVVVPNRHLIY-----
NRTF1c_ -----SLAELLNKLAK-----NRSQKV E C-----
NRTF1d_ -----SLAELLNKLAK-----NRSQKV E C-----

```

**Figure 5:** Amino acid alignment of soybean NRP-1 (SEQ ID NO: 15) and NRP-2 (SEQ ID NO: 17) along with a tomato miraculin homologue (T07871, SEQ ID NO: 25) and tobacco tumor-related protein (T03803, SEQ ID NO: 26)

```

T03803      M K T N Q L F L P F L I F T I S F N S F L S S S A E A P P A - V V D I A G K K L R T G I D Y Y I L P - - - - V V R G - -
T07871      M K I N Q L F F P F L I L A I S F N S L L S S A A E S P P E - V V D I D G K I L R T G V D Y Y I L P - - - - V V R G - -
NRP-1       - M K T K L L A F L L F F A L T T K P L L L G A A G A A P E P V I D T S G K K L R A D A N Y H I I P A V P F T I C G F V
NRP-2       M K S T M L L A F A L V L A L S S Q P - L L G G A E A S P E Q V V D T L G K K L R V G T N Y Y I V P S L P Y T K I R - -
              . * :   * . : : : : : . : * . . * : . *   * : *   * * * . . : * : * : *   .

T03803      - - - R G G G L T L D S T G N E S C P L D A V V Q E Q Q E I K N G L P L T F T P V N P K K G V I R E S T D L N I K F S -
T07871      - - - R G G G L T M D S I G D K M C P L D A V V Q E H N E I D Q G L P L T F T P V D P K K G V I R E S T D L N I I F S -
NRP-1       S C F T G G G L S L D S I D E S - C P L D V I I E K A N - - - E G L P L R F S P V N T K K G V I R V S T D L N I F F S D
NRP-2       - - - T T R G L G L A S V G K P Y C P L D V V V V N G Y - - - H G L P V T F S P V N P K K G V I R V S T D L N I K F S -
              * * : *   . .   * * * . : : :   :   . * * * : * : * : . * * * * * * * * * * * * * *

T03803      A A S I C V - - Q T T L W K L D D F D E T T G K Y F I T I G G N E G N P G R E T I S N W F K I E K F E R D Y K L V Y C P
T07871      A N S I C V - - Q T T Q W K L D D F D E T T G Q Y F I T L G G D Q G N P G V E T I S N W F K I E K Y D R D Y K L L Y C P
NRP-1       S D E R C P - H H S T V W M L D Q F D A S I G Q T Y V T T G G V V G N P G E H T I L N W F K I Q K Y E D A Y K L V Y C P
NRP-2       A R T S C P R Q Y S T V W K L D D F D F S K R Q W F V T T G G V V G N P S L E T I H N W F K I E K Y D G A Y K L V Y C P
              :   *           : * *   * * : * * :   :   : : * *   * * * . . * *   * * * * : * :   * * * : * * *

T03803      T V C N F C K V I C K D V G I F I Q D G I R - R L A L S D V P F K V M F K K A Q V V K D
T07871      T V C D F C K V I C R D I G I F I Q D G V R - R L A L S D V P F K V M F K K A - - - - -
NRP-1       R V C P S C H H L C K D I G M F V D A N R R M H L A L S D D P F K I K F K E A - - - - -
NRP-2       S V V K C P K H L C K N V G L F V D E K G N K R L A L T D V P L K V Q F Q Q A - - - - -
              *           : : * : : : * : : :   .   : * * * : *   * : * :   * : *   .

```

Figure 6: Amino acid sequence alignment of soybean 7OM sequence (SEQ ID NO: 19) compared to maize 7OM (L14063, SEQ ID NO: 27) and *Medicago* 7OM (AF000975, SEQ ID NO: 28)

AF000975	MASSINGRKPSSEIFKAQALLYKHIYAFIDMSLKWAVEMNIPNIIQNHGKPISLSNLVSI
7OM	MASMNN-QKEIELFEGQSLLYMQLYGHLRPMCLKWAVQLGIPDIIQNHAKPISLSDLVST
L14063	MELSPNNSTDQSLLDAQLELWHTTFAFMKSMALKSAIHLRIADAIHLHGGAASLSQILSK
	* * . . . . . * * : : : . * . * * * : : : * : * : * . . * * : : *
AF000975	LQVPSSKIGNVRRLMRYLAHNGFFEIITK----EE---ESYALTVASELLVRG--SDLCL
7OM	LQIPPANAAFVQRFMRFLAHNGIFEIHES----QEDHELTALTPASKLLVNS--SDHCL
L14063	VHLHPSRVSSLRLMRVLTNTNVTGTPQLGGGSDDDSEPVYTLTPVSRLLIGSQSSQLAQ
	: : . . . : * : * * * : : * : : : : * : * * * : : * : .
AF000975	APMVECVLDPTLSGSYHELKKWIYEE--DLTLFGVTLGSGFWDFLDKNPEYNTSFNDAMA
7OM	SPMVLAFTDPLRNVKYHHLGEWIRGE--DPSVFETAHGTSAWGLLEKNPEYFSLFNEAMA
L14063	TPLAAMVLDPTIVSPFSELGAWFQHELDPDCIFKHTHGRGIWELTKDDATFDALVNDGLA
	: * : . . * * : . * * : * * : * : * : . . . : : : * : : * *
AF000975	SDS-KLINLALRDCDFVFDGLESIVDVGGGTGTTAKIICETFPKLKCIIVDRPQVVENLS
7OM	SDS-RIVDLALKNCTSVFEGGLDSMVDVGGGTGTTARIICDAFPKLKCVVLDLPHVVENLT
L14063	SDSQLIVDVAIKQSAEVFQGISSLVDVGGGIGAAAQAISKAFPHVKCSVLDLAHVVAKAP
	* * * : : : * : : : . * * : : . * : * * * * * * : : : * : : * * * * * : .
AF000975	GSNNLTYYVGDMFTSIPNADAVLLKYILHNWTDKDCRLRILKKCKEAVTNDGKRGKVTIID
7OM	GTNNLSFVGGDMFNSIPQADAVLLKWVLHNWTDENCIKILQKCRDSISSKGN SGKVIIID
L14063	THTDVQFIAGDMFESIPPADAVLLKSVLHDWDHDDCVKILKNCKKAIPPREAGGKVIIN
	: : : : * * * * * * * * * * : * : * : : * : : * : : : : . * * * * :
AF000975	MVIDEKKDENQVTQIKLLMDVNMACLNGKERNEEEWKKLFIEAGFQHYKISPLTGFLSLI
7OM	AVINEKLDDPDMTQTKLSLDIIMLTMNGRERTEKEWKQLFIEAGFKHYKIFPIFGFRSLI
L14063	MVVGAGPSDMKHKEMQAFDVYIMFINGMERDEQEWSKIFSEAGYSYRIIPVLGVRSLI
	* : . . . : : : * : : : * * * * * : * * * : : * * * : * : * * *
AF000975	EIYP----
7OM	EVYPXTFL
L14063	EVYP----
	* : *

Figure 7. Amino acid sequence alignment of soybean and *Arabidopsis* IPP proteins (SEQ ID NOs: 21 and 29, respectively).

```

AY048296      MSPVEPAGIMKKSHRQKSQRLWAKLVMRKWLNISGRDPEYGADTDNESENEDEDAREDNDDS
IPP           -----MWPTLVANKIFKKRLGSSNFIADYPSYKEPLLGVVDID--
                        :*. ** . * ::      .::: ** . . * . * *

AY048296      SSDEEGSGSRGRESKVYENAEDAIAAASAVVDAAAAAAEFISNDAPMKLRRRNSETLRA
IPP           -----QNSKTILN
                        :*: *:

AY048296      QYINNKEIRVCVGTWNVGGISPPSDLDIDDWIEIN-QPADIYVLGSQEIVPLNAGNILGA
IPP           DHKDSHKYKVFVSTWNVGGIAPDEDLNIDDLLETENNCDIYILGFQEIVPLKASNVLGS
                        :: :.::: : * *.*****: * .*:*** : *      :..***: * *****:*. :*:
                        :. :.::: : * *.*****: * .*:*** : *      :..***: * *****:*. :*:

AY048296      EDRPVPKWEVIREALNVRVRPKLSGVKSYSDPPSPGRFKPFEETHDIIIEEVAFESDSD
IPP           ENNEISMKWSNIREALN-----KKITHQR-DK-
                        *:..      **:.:*****
                        :. :. . * .

AY048296      AGVEIHPIDEEEEETDRLWALKHDGGVIGEVKTLVDPNTGLPVVEIKRQFSIPKKLDRQ
IPP           -----DAKKQE-----LKN-----F-PN-----K-----
                        * :.: *      **: :      . **      :

AY048296      LCLRADSFKGISDDSTQTGMKTINRLSGKERIGLSWPEPPLNMLGPCVLDRQPSIKTV
IPP           -----KE-----NP-----
                        **
                        : *

AY048296      KSLKTAKSFKAYSSFKSVAGNNNGIPPEVLALAEMDLKLLMERKRRPAYVRLVSKQMVGI
IPP           ----AKCCDAPHDFQCIIS-----KQMVGL
                        ** . * . *:.: .
                        *****:

AY048296      LLTIWVKRSLRKHIQNVRVSTVGVMGYIGNKGAVSVSMSINQTFFCFINTHLAGERE
IPP           FISVWIRRDLCPIRHPSVSCVCGCIMGCLGNKGSISVRFLHETSFCFVCSHLASGGRE
                        :::::*. * .*: : ** * * *: * :*:*: * :.: * * * :*: * *

AY048296      VDQIKRNADVHEIHKRTVFHVSALGLPKLIYDHERIIWLGDNLNRLSSSYEKTRDLISK
IPP           GDEKHRNSNVAEIFSRTSFPRGPLLDLPRTILDHDHVLGDLNRYRISLPEETRLLEK
                        *: :*: : * *. ** * . * .*: : * * *: * *****: * . * . * *:
                        :. :.::: : * *.*****: * .*:*** : *      :..***: * *****:*. :*:

AY048296      REWSKLEYPDQLVKEYRKGRAFDGWSEGTLHFPPTYKYQANSDEYTAND---GKAPKR-T
IPP           RDWDSLLANDQLIMELMSGNMLRGWHEGAIKFAPTYKYCPNSDIYGYCCYHGKKAEKRRR
                        *:*. ** *****: * . * . : ** * *: :*. ***** .*** * .      ** * :

AY048296      PAWCDRVLSYGKGMRLVHYRRTEQKFSDRPVTAIYMAEVEVF-----SARKL
IPP           PAWCDRIVWCGEGLKQLQYTRIESKLSDRPVKAMFIAEVRVLPPELMKNLQSLFLSERYE
                        *****: :*: : : * * *.*:*****. :*:***. *:
                        * *

AY048296      QRALTFTDAEIEDEG-----LVAVLV-----
IPP           QIKTPFEVSTTEDFVNRKRSSFRLXIFCVIQRAL
                        * . * : **      * . *

```

Figure 8: 5' Flanking Region of the soybean IPP gene (SEQ ID NO: 22) showing the TATA box and first ATG codon

AGCACCATCATCTCTATCATTTCGGAATGCAACCAAGCTAAAAGATTACTACAACCTAATTGCTTTTCCTTATCACA  
TTTTCTGACGTATAGTGTGATTTTATATATTTTACATAAGAGAAAATAATTTCTACAAAAAAATGTTTCGATAAA  
ATAAGTAGAGAACGTGTATTAATAATTTCTACATAAGAAAATAAGAAATATATTAGATATAATAAGTGATGCAAG  
AAAGATGGACAAAAATAATTACGTAAATATCATTCTATAAAATTTATTATTTCATTATATAAATAGCATTACCATTG  
TTGAAACTTGAAAGTGGGTCCATCGTTTACAACCTAAAGAAAAGACACCCTAGCGTAAAATATTCAACCATCGACGT  
CTACTTCAATTAACATGAAGATGTAGTTCCATCTCAACGGATTTCCGTCTCAAATAAAATTCCTTAATAACGTGCT  
ACTAACCATTGGAATCTGCAGAATATCTCGTTTAGTTGGGCACAATCCCTCAAAGCGATGTATTTTTTTAATGG  
AAACAATGCATGCCACAAGAACGTTTATATATACATAATTTTACTAAACAAATCGTAATACAAAACCTTATTATT  
ATAACGTGATTTGTCACCTTTTGTCTTCAGAAAAATCTTTGTACAAAACATTAAGACAATAACATAAGTTGCCAA  
TACCATACATAAAAACCTTTTAATGAATCATAATGATGAAAAATTGAGAGATATTTAGTTCCATGATAAAGAGTGTG  
TTTGTGTGGGAATTTGACCAAACGCAATTGTTGTTCCAGTGAAAACTTTTCTCGCGTGTTTGGCCTTTTGTGTCT  
CAGAAAGCTAATTTTCTCCATTTAACGTGGTTTGGACCCATTTTCAAACGCACTCACAGTGAGTCCGTTTCTGTA  
GAAGTCCTTCCAGCAGGCCACCCCTCCCCCAAAGAAAAATTTAAGAAACAAGAAGAGAAAAGAAAAAAGAAAAAG  
AAAAAACAAGTCACATATTTTATTCTTATGTGAGCCAAAAACCTTGACTAGCTGTAGATGGGGCAATAATAACTAG  
CTATTCATCACATTTCTAGCTAATTGCCTGTTTGTATGGACCACATTCCCCTTGCACTCATCTTCAGCAAT  
TTAAATTAGGTAATAAACATTAAGATATCCTTTAAAATCTACTCAACAAACAGAAGAATCAAATCTGCAAGAAG  
GGTAGACCCATGTTTATACTAGCTTTCTCTATCTCTCTCCCACTGGGACATAAATGTTCCATATTTCAAAAAA  
TATATATATGATATGATGAGCAATGCAGCCAAAGGTGCATCATCTTTTACGTCACATGAAAGCCTTCTACCTC  
TTCAAGCTGCACAAGCCTTTCTCTTTCCAGAAATGATTTTTTCCATTTCTTGTATTATTACTCCTTTTGGACT  
TTCTATATAATGCTTTCTATATACGTTTCCAATAATACCACGTACACTACTCATGTGCCAGGAAAAGGAGCAGCA  
GTGACCACCTAGCAATAGTACTCTCGCCTTCTCTCAATCATTTTTTCATTTGTCAACTTTTATAGACCTCGATTTG  
TGTATGCAAAATTTCTTTAGGACATG